Having thus described our invention, what we claim as new, and desire to secure by Letters Patent is:

1. A high speed DRAM, comprising:

a DRAM memory;

a cache memory;

a read register coupled between the cache memory and the DRAM memory, for transferring data from the cache memory to the DRAM memory;

a write register coupled between the DRAM memory and the cache memory, for transferring data from the DRAM memory to the cache memory;

a first bi-directional data bus set coupled between the cache memory and both the read register and the write register;

a second data bus set coupled between the read register and the DRAM memory;

a third data bus set coupled between the DRAM memory and the write register.

2. The high speed DRAM of claim 1, wherein the cache memory comprises a single port SRAM.

3. The high speed DRAM of claim 1, wherein a multiplexer couples the cache memory to either of the read register or the write register, and a fourth data bus couples the multiplexer to the read register, and a fifth data bus couples the multiplexer to the write register.

4. The high speed DRAM of claim 1, wherein a sixth data bus couples the read register to a data output from the circuit, and a seventh data bus couples a data input to the circuit to the write register.

5. The high speed DRAM of claim 4, wherein an eighth data bus couples the write register to outside data buses.

6. The high speed DRAM of claim 5, wherein a multiplexer switches between inputs received from the sixth data bus from the read register and the eight data bus from the write register, and outputs data onto a ninth data bus coupled to the outside data buses.

7. The high speed DRAM of claim 6, wherein a read buffer couples the read register to the DRAM memory through a tenth data bus.

8. The high speed DRAM of claim 7, wherein an eleventh data bus couples the DRAM memory to the write buffer which is coupled through the third data bus to the write register.

9. The high speed DRAM of claim 7, wherein the first, second, third, fourth, fifth, tenth, and eleventh data buses all have the same first wide data bandwidth.

10. The high speed DRAM of claim 6, wherein the sixth, seventh, eight, and ninth data buses all have the same second narrow data bandwidth;

11. The high speed DRAM of claim 1, wherein data flows through the bi-directional bus in a first direction from the cache memory to the read register, and data flows through the bi-directional bus in a second opposite direction from the write register to the cache memory, such that opposite direction data flows share the same bi-directional data bus in different cycles.

12. The high speed DRAM of claim 1, wherein data flows from the read register to the DRAM memory in a first cycle, and data flows from the DRAM memory to the write register in a second cycle, to share access to the DRAM memory in different cycles.

13. A method of operating the high speed DRAM of claim 1, wherein for a read miss operation, a new set of data are retrieved from the DRAM memory to replace old data in the cache memory, and also to be sent to outside data buses through an output read buffer, and during a first cycle of data flow, data flows from the cache memory through the first and fourth buses and is latched into the read register, and data coming from the DRAM memory are latched into the write register through the third bus, and in a second cycle, the directional flows of the data are reversed through the first and fourth buses and also through the third bus.

14. A method of operating the high speed DRAM of claim 1, wherein for a write miss operation, a new set of data are written into the cache memory to replace retired data, partly from outside data buses via a write buffer, and the rest of the data are from the DRAM memory, and these data are merged in the write register.

15. A method of operating the high speed DRAM of claim 1, wherein for a read hit operation, data are transferred nondestructively from the cache memory through the read register to an output read buffer via a multiplexer, and according to a column address, only a portion of the data are transferred to outside data buses.

16. A method of operating the high speed DRAM of claim 1, wherein for a write hit operation, a new set of data is transferred to the cache and overwrite a portion of the old data therein.

17. A method of operating the high speed DRAM of claim 1, wherein for a two cycle read hit operation, data that resided in the cache memory are read out according to row address and are latched into the read register based upon column address, and only a portion of these data are transferred to outside data buses via the multiplexer and an output read buffer, and

in a first clock cycle, data are latched in sense amplifiers of the cache memory, and

in a second clock cycle, data are latched and decoded in the read register.

18. A method of operating the high speed DRAM of claim 1, wherein for a two cycle write hit operation, upon detecting a write address is in the cache memory, data is transferred from outside data buses to the cache memory, these data flow via an output write buffer and are then latched into the write register and only occupy a portion of the write register, and only this portion is written into the cache memory based upon column address, and the rest of the data in the same row of the cache memory is maintained unchanged,

in a first clock cycle, data are written into the write register, and

in a second clock cycle these data are latched into sense amplifiers of the cache memory.

19. A method of operating the high speed DRAM of claim 1, wherein for a three cycle read miss operation, upon detecting that read data are not resident in the cache memory, then old data with the same row address are written back into the DRAM memory in a fast cycle DRAM operation wherein original data are destroyed after they are read into the cache memory, and therefore when these data are not needed in the cache memory, they are written back to the DRAM memory to prevent a data loss.

20. A method of operating the high speed DRAM of claim 1, wherein for a three cycle write miss operation, upon detecting that write data address is not resident in the cache memory, then old data in the same row of the cache memory are written back into the DRAM memory, and

in a first cycle, old data are latched in sense amplifiers in the cache memory while new data are latched in DRAM memory sense amplifiers and a set of the new data are latched into the write register, and

in a second cycle, old data are transferred to the read register, and at the same time new data are transferred from the DRAM memory to the write register, wherein based on the column address, data from the DRAM memory and a set of data from outside data buses are merged,

in a third cycle, old data are transferred and latched into the DRAM memory while new data are sent to the cache memory.

21. A method of operating the high speed DRAM of claim 1, wherein for a write back operation, which is needed for both a read miss operation and a write miss operation while old data are written back to the DRAM memory, a new set of data from the DRAM memory with a correct row address is read into the write register and then to the cache memory to replace the old data, while retrieving these data, a portion of the data are read to outside data buses based upon column address, decoding is performed in the write register, a selected set of data are transferred to the outside data buses via an output read buffer, wherein two streams of data are transferred simultaneously in two opposite paths via two sets of bus sets,

in a first clock cycle, old data are latched into the cache memory sense amplifiers, while new data are latched in the DRAM memory sense amplifiers,

in a second cycle, old data are latched into the read register while new data are latched into the write register, and at the same time a set of the data are sent to the outside data buses and are latched into a read buffer,

in a third cycle, old data are written back into the DRAM memory, and new data from the DRAM memory are transferred into the cache memory to replace old data.